UNITED STATES PATENT APPLICATION

for

CONTENT AND DISPLAY DEVICE DEPENDENT CREATION OF SMALLER REPRESENTATIONS OF IMAGES

Inventors:

Kathrin Berkner

Ramesh Neelamani

Gregory J. Wolff

Martin Boliek

Peter Hart

prepared by:

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP

12400 Wilshire Boulevard

Los Angeles, CA 90025-1026

(408) 720-8300

File No.: 074451.P137

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number: __EL 617 207 575 US__

Date of Deposit: __January 10, 2002__

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Box Patent Application, Assistant Commissioner for Patents, Washington, D. C. 20231

__MELANIE LYONS__

(Typed or printed name of person mailing paper or fee)

(Signature of person mailing paper or fee)

_1-10-02_

(Date signed)

CONTENT AND DISPLAY DEVICE DEPENDENT CREATION OF SMALLER REPRESENTATIONS OF IMAGES

RELATED APPLICATIONS

[0001]   This application is related to the co-pending application entitled *Header-Based Processing Of Images Compressed Using Multi-Scale Transforms,* concurrently filed on January 10, 2002, U.S. Patent Application Serial No. _____, assigned to the corporate assignee of the present invention.

FIELD OF THE INVENTION

[0002]   The invention relates generally to the field of image processing. More specifically, the invention relates to creating a smaller representation of an image.

BACKGROUND OF THE INVENTION

[0003]   Thumbnails are an extremely desirable graphical user interface component for most multimedia and document applications. A thumbnail is a resized smaller version of an image representative of the full image which can be displayed in some applications by clicking on the thumbnail. The resizing is typically done by traditional smoothing followed by downsampling. In most traditional applications, such as listings on a web page, the size of a thumbnail is fixed. A common problem with those thumbnail displays is that the image information is often not recognizable for the viewer and does not provide the desired usefulness.

[0004]   Newer multimedia communication tools allow a free-format composition of images of various sources on a representative canvas. In this case, the size of a

thumbnail is allowed to be variable. In such an application, besides the question of what to display in a fixed-size thumbnail, the additional question arises of what constitutes a suitable thumbnail size or shape for a given image. For example, in representing a photo of a person as captured from a visitor's kiosk, a downsampled image of the face-portion of the photo would be sufficient for a thumbnail, whereas for a web document the title at full resolution might be a useful display.

[0005]   The authors in Mohan, R., Smith, J.R., and Li, C.-S., "Adapting Multimedia Internet Content for Universal Access," IEEE Trans. Multimedia, Vol. 1, no. 1, pp. 104-114, 1999, describe a method for the transcoding of images to specific devices such as monitors or cell phones. In Lee, K., Chang, H.S., Choi, H., and Sull, S., "Perception-based image transcoding for universal multimedia access," which appeared in ICIP 2001, Proceedings of International Conference on Image Processing, Thessalonihi, Greece, 2001. This approach is extended to sending only specifically selected parts of an image at a specific resolution. The specification is performed by the sender and does not happen automatically.

[0006]   There exist several software packages providing for thumbnail creation. These software packages focus on speed, but all of these resize the entire image to user or application defined sizes using traditional downsampling. Thus, image information is often not recognizable to the viewer.

[0007]   In Woodruff, A., Faulring, A., Rosenholtz, R., Morrison, J., and Pirolli, P., "Using Thumbnails to Search the Web," Proceedings of SIGCHI'01, Seattle, April 2001, enhanced thumbnails are introduced to provide a better representation of documents. The enhancement consists of lowering the contrast in traditionally created thumbnails and

superimposing keywords in larger fonts that were detected via an Optical Character Recognition (OCR) system. The result is a limited improvement at best and is only applicable to images that contain text.

[0008]   The authors in Burton, C.A., Johnston, L.J., and Sonenberg, E.A., "Case study: an empirical investigation of thumbnail image recognition", Proceedings of Visualization Conference 1995, found that the filtering of images (contrast enhancement, edge enhancement) before downsampling increases a viewer's ability to recognize thumbnails. Even so, image information is often not recognizable to the viewer.

[0009]   For creation of video summaries there exist methods that display groupings of video frames where the individual frames have specific sizes/resolution, as in Uchihashi, S., Foote, J., Girgensohn, A., and Boreczky, J., "Video Manga: Generating Semantically Meaningful Video Summaries," Proceedings of Seventh ACM International Multimedia Conference, Orlando 1999. The decision for resizing the frames is made by measuring the frames importance in the video sequence, not by the actual image content.

[0010]   The area of page layout has been discussed in many papers. Typically, the authors assume, disadvantageously, that the image content of documents is mostly text with perhaps some small images or graphics, and perform text specific operations involving clustering techniques to determine connected components. A common method for page layout is the one described in O'Gorman, L., "The Document Spectrum for Page Layout Analysis," IEEE Trans. Image Proc., Vol. 15, no. 11, pp. 1162-1173, 1993.

[0011]   One existing image file format stores multiple resolutions of an image (e.g., as created by a Laplacian pyramid). As a result, this image file format is usually disadvantageously larger than a file of a wavelet coded image. This image file format has

the option of incorporating specific parameters into the file. Those parameters could be, for example, result aspect ratio, rectangle of interest, filtering or contrast adjustment. The parameters are set by the creator of the file and may or may not be used by the receiver.

SUMMARY OF THE INVENTION

[0012] A method and apparatus to receive an image and to create a smaller representation of the image from a wavelet representation of the image is described. The form of display (e.g., size, shape, multiscale collage) of the smaller representation of the image is selected to compensate for the content of the image and physical properties of the display device to display the smaller representation of the image.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013]   The present invention will be understood more fully from the detailed description given below and from the accompanying drawings of various embodiments of the invention, which, however, should not be taken to limit the invention to the specific embodiments, but are for explanation and understanding only.

[0014]   Figure 1 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0015]   Figure 2 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0016]   Figure 3 illustrates a typical wavelet decomposition;

[0017]   Figure 4 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0018]   Figure 5 illustrates multiple sampling factors (display scales);

[0019]   Figure 6 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0020]   Figure 7 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0021]   Figure 8 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0022]   Figure 9 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0023]   Figure 10 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0024]   Figure 11 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0025]   Figure 12 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0026]   Figure 13 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0027]   Figure 14 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0028]   Figure 15 is a flow diagram showing results of an approach for computing an image representation that is visually recognizable when being displayed on a given display device for two example images;

[0029]   Figure 16 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0030]   Figure 17 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0031]   Figure 18 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0032]   Figure 19 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given device.

[0033]   Figure 20 is a flow diagram of one embodiment of a process for image creation based on global scale selection;

[0034]   Figure 21 is a flow diagram of one embodiment of a process for image creation based on local scale selection;

[0035]   Figure 22 is a flow diagram of one embodiment of a process for image creation based on a combination of global and local scale selection;

[0036]   Figure 23 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0037]   Figure 24 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0038]   Figure 25 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device;

[0039]   Figures 26A-C illustrates exemplary design choices;

[0040]   Figure 27 illustrates additional exemplary design choices;

[0041]   Figure 28 is a block diagram of one embodiment of a device for computing an image representation that is visually recognizable when being displayed on a given display device; and

[0042]   Figure 29 is a block diagram of an exemplary computer system.

DETAILED DESCRIPTION

[0043]   A method and apparatus for automatic creation of smaller-sized representations of images is described that takes into account the image content as well as the physical properties of a display device, such as, for example, but not limited to, dots per inch (dpi) resolution, absolute pixel resolution, pixels per viewing angle, gain function (gamma response), contrast, and brightness, gain function to compute an image representation that contains information that is visually recognizable when being displayed on the specific display device. The techniques described herein use a wavelet representation of an image and parameters that characterize the image content.

[0044]   The techniques described herein also use parameters linked to physical properties of a display device. In a typical Internet application, an image may be displayed on several different display devices such as a monitor, a media board, a digital camera display, a personal digital assistant, or a cell phone. These devices may differ in physical properties such as, for example, dots per inch, numbers of pixels, contrast, brightness, etc.

[0045]   In the following description, numerous details are set forth to provide a thorough understanding of the present invention. It will be apparent, however, to one skilled in the art, that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the present invention.

[0046]   Some portions of the detailed descriptions which follow are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means

used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

[0047]   It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussion, it is appreciated that throughout the description, discussions utilizing terms such as "processing" or "computing" or "calculating" or "determining" or "displaying" or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

[0048]   The present invention also relates to apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but is not limited to, any type of disk

including floppy disks, optical disks, CD-ROMs, and magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus.

[0049]   The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present invention is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the invention as described herein.

[0050]   A machine-readable medium includes any mechanism for storing or transmitting information in a form readable by a machine (e.g., a computer). For example, a machine-readable medium includes read only memory ("ROM"); random access memory ("RAM"); magnetic disk storage media; optical storage media; flash memory devices; electrical, optical, acoustical or other form of propagated signals (e.g., carrier waves, infrared signals, digital signals, etc.); etc.

OVERVIEW

[0051]   Figure 1 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. This process, like others described below, is performed by processing logic that may comprise hardware, software, or a combination of both.

11

[0052]   Referring to Figure 1, processing logic initially receives an image (processing block 101). Processing logic then creates a smaller representation of the image from a wavelet representation of the image (processing block 102). The size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device that is to display the smaller representation of the image.

[0053]   Figure 2 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 2, processing logic receives an image (processing block 201). Processing logic then creates a smaller representation of the image from a wavelet representation of the image, where the size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 202). In contrast to Figure 1, the physical properties include dots per inch (dpi) resolution, absolute pixel resolution, viewing distance, contrast, and brightness.

DETERMINATION OF A DISPLAY DEVICE DEPENDENT RANGE OF DOWNSAMPLING FACTORS

[0054]   A wavelet representation separates the image information into its information at various levels of resolution 1...L. Figure 3 illustrates a typical wavelet decomposition. If an image 301 of size NxM is given in its wavelet decomposition as if to be downsampled by, for example, a factor of 4 in each dimension, then the inverse wavelet transform is performed on levels 3...L. The result after this inverse transform is the LL component at level 2 (i.e., a downsampled version of the original image). The size of the downsampled image is $M/(2^2) \times N/(2^2)$. This reduced image will not contain

12

the detail information contained in the highpass coefficients at level 1 and 2 of the wavelet decomposition.

[0055]   Depending on display device parameters, a range of suitable display levels between level 1 and L can be determined as follows. In one embodiment, if the size of the image is larger than the pixel size of the display device (e.g., the image is of size 2048x2048 and the monitor resolution is of size 1024x768), then the image should be downsampled at least twice to be completely visible on the monitor display. Therefore, the absolute display resolution provides a lower bound on the number of downsampling factors (i.e., on the display level L, so called L_{min}).

[0056]   Figure 4 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 4, processing logic receives an image (processing block 401). Processing logic then creates a smaller representation of the image from a wavelet representation of the image, where the size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 402). As shown in Figure 4, as part of creating the smaller representation, processing logic downsamples the image a number of times (processing block 403). The number of times is large enough to cause the smaller representation of the image to be completely visible on the display device. For example, if an image has 2048 x 2046 pixels and the monitor has a 1024 x 768 pixel display resolution and downsampling is performed once, then the image will not fit in the display area of the monitor; however, if downsampling is performed twice, then the image will fit in the display area of the monitor.

13

[0057]   Another factor that influences the visibility of images on displays is the relative display resolution given in dots per inch (dpi). For example, a monitor has a relative resolution of 75 dpi, whereas a printer has a resolution of 600 dpi and a large display has an estimated resolution of 25 dpi. Other tests may have other results (e.g., text recognition). As determined from subjective tests, the minimal size of a visible recognizable object or structure is a diameter of approximately 2mm. This size can be translated into a dot size for a given dpi resolution. For a 75 dpi monitor, the minimal object diameter in dot numbers would be 6, whereas for a 600dpi printer, it would be 47 dots. In one embodiment, the formula for determination of minimal visible object diameter in dot numbers is given by

$$D_{min} = (\text{minimal object diameter in mm})*(\text{dpi resolution})/25.4.$$

[0058]   Depending on physical properties of the display device, such as contrast and brightness, the choice of D_{min} may vary a bit. In one embodiment, this variation is expressed as a constant C_{device} in the formula and tuned for a specific display. In one embodiment, if contrast or brightness is small, then the constant C_{device} is large.

$$D_{min} = C_{device} * (\text{minimal visible object diameter in mm})*(\text{dpi resolution})/25.4. \quad (1)$$

D_{min} also depends on the viewing distance and angle. In one embodiment, the ratio between dpi resolution and view distance is constant.

[0059]   Small scale structures show up as wavelet coefficients in decomposition levels over a number of decomposition levels associated with high resolution scales. In one embodiment, to be visibly recognizable after n-times downsampling, small scale structures have a dot size D that is larger than D_{min} after n-times downsampling. Therefore, the inequality $D/2^n > D_{min}$ is valid. As a result, only structures that correspond

14

to wavelet coefficients being nonzero at levels n...L qualify to be visibly recognizable at the chosen display device. Thus, the levels 1...n-1 do not need to be considered at all during the inverse transform, since they will not provide visually recognizable information. The level

$$L_{max}(D) = \log_2(D/D_{min}) = \log_2(D) - \log_2(D_{min}) \qquad (1)$$

provides an upper bound on the downsampling factor of an object of dot size diameter D. As an immediate consequence,

$$\lfloor L_{max} \rfloor = \lfloor \log_2(\min(N,M)/D_{min}) \rfloor = \lfloor \log_2(\min(N,M)) - \log_2(D_{min}) \rfloor \quad (2)$$

is an upper bound on the downsampling factor for an entire image of size NxM. L_{max} and L_{min} form a range of possible display scales

$$\lambda = [L_{min}, L_{max}]. \qquad (3)$$

The bound L_{max} is most likely too large for most images. Often, images contain almost equal proportion of coarse and fine structures, such as natural images, or a lot of fine structures, such as text images. Depending on the ratio of coarse to fine structures, a natural image may be downsampled more times than a text image and still display visually recognizable information. Therefore, in one embodiment, the information contained in wavelet coefficients is used to determine, for a given image, the maximal downsampling factor $L_{display}$ under the constraints that

$$L_{min} \leq L_{display} \leq L_{max} \qquad (4)$$

and that the LL component at level $L_{display}$ contains important image information. Figure 5 illustrates an overview of sampling factors (display scales). The factors are shown along a line of all possible downsampling factors 501. The range bounds L_{min} 502 and L_{max} 503 depend on display parameters, dpi resolution, absolute pixel resolution,

15

brightness and contrast, viewing distance, and the image parameter image size. In addition to those parameters, $L_{display}$ 504 depends also on the image content. Another factor, $\log_2$(image size) 505, is also present at one extreme of the line of all possible downsampling factors 501.

[0060]   Figure 6 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 6, processing logic receives an image (processing block 601). Processing logic then creates a smaller representation of the image from a wavelet representation of the image, where the size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 602). In this case, the size of the smaller representation of the image depends on a minimal visible object diameter of an object in the image. The minimal visible object diameter is at least a number of dots. The number of dots depends proportionately on a dots per inch resolution of the display device.

[0061]   Figure 7 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 7, processing logic receives an image (processing block 701). Processing logic then creates a smaller representation of the image from a wavelet representation of the image, where the size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 702). In this case, the size of the smaller representation of the image depends on a minimal visible

16

object diameter of an object in the image. The minimal visible object diameter is at least

a number of dots. The number of dots depends proportionately on a dots per inch

resolution of the display device. As shown in Figure 7, as part of creating the smaller

representation, processing logic downsamples the image a number of times (processing

block 703). The number of times is small enough to cause a number of dots in a diameter

of the object to be at least as large as a number of dots in the minimal visible object

diameter. For example, in the case where the minimal object diameter is 20 dots and the

object diameter is 100 dots, then downsampling twice results in a diameter of 25 dots.

Downsampling three times results in a diameter that is 12 dots. Since the 12 dots is less

than 20 dots, the number of times the object can be downsampled is only two.

[0062]   Figure 8 is a flow diagram of one embodiment of a process for computing

an image representation that is visually recognizable when being displayed on a given

display device. Referring to Figure 8, processing logic receives an image (processing

block 801). Processing logic then creates a smaller representation of the image from a

wavelet representation of the image, where the size of the smaller representation of the

image is selected based on the content of the image and physical properties of a display

device to display the smaller representation of the image (processing block 802). As

shown in Figure 8, as part of creating the smaller representation, processing logic

downsamples the image a number of times (processing block 803). The number of times

depends proportionately on a ratio of coarse structures in the image to fine structures in

the image.

[0063]   The display scale $L_{display}$ is computed from the distribution of wavelet coefficients over the device dependent display range $[L_{min}, L_{max}]$. In one embodiment, large coefficients correspond to important edges in the image and small coefficients correspond to unimportant small-scale structures and noise. In one embodiment, a level in the decomposition is determined such that the LL component at that level contains the most important information to allow for visual recognition of the image. Two approaches are described below.

[0064]   In one embodiment, the best display scale is determined for the entire image from certain resolution levels (i.e., global scale selection). In an alternative embodiment, the image (or its wavelet decomposition) is partitioned into cells and the best display scale for each cell is determined independently (i.e., local scale selection). These approaches are further described below.

GLOBAL SCALE SELECTION

[0065]   In one embodiment, a mathematical formulation for the global scale selection problem is as follows. Consider a wavelet decomposition of an image of size MxN. $W_m^{LL}(X)$ denotes the set of all LL coefficients of the image X at level m. Similarly, $W_m^{HL}(X)$, $W_m^{LH}(X)$, $W_m^{HH}(X)$ denote the set of all wavelet coefficients in the HL, LH and HH bands at level m.

[0066]   M denotes a measure that reflects the importance of information contained in the low resolution version $X_m = W_m^{LL}(X)$. This importance measure may be defined

based on one or more factors, such as, for example, but not limited to, energy, entropy, etc.

[0067]   The best display level $L_{display}$ is computed by

$$L_{display} = \arg \quad max_{m\epsilon\lambda} \quad \{Cost(M(W_m^{LL}(X)))\}. \tag{5}$$

The actual resulting image is then defined as $W_{Ldisplay}^{LL}(X) = X_{Ldisplay}$.

[0068]   There are a number of different approaches. Two different approaches include scale selection by maximal energy or entropy and scale selection by two-class labeling.

SCALE SELECTION BY MAXIMAL ENERGY/ENTROPY

[0069]   In one embodiment, the display scale is determined by characterizing trends of wavelet coefficients over the levels of resolution and detect levels of trend change. One criterion for a trend is the energy in subbands. If $E_{LH}[m]$ denotes the energy of wavelet coefficients in the subband LH at level m for a text image, then the trend is typically $E_{LH}[L]<E_{LH}[L-1]<...<E_{LH}[1]$. The same trend is valid for the other two subband directions, HL and HH. The trend for a photographic natural image has the opposite behavior, $E_{LH}[L]>E_{LH}[L-1]>...>E_{LH}[1]$. An image that contains mainly text in a small font will follow the text trend for all scales. An image with medium size text will follow the text trend for some large scales $m_1...L$, and then will switch to the photo trend for the small scales $1...m_1-1$. Therefore, the smaller scales contain less information than the larger scales and downsampling to scale m_1 still keeps most of the important information in the image. The display scale is then chosen to be m_1. In one embodiment,

19

considering the energy as a function of level E(m), the scale at which the global

maximum is located [L_{min},L_{max}] is determined by

$$L_{display} = \arg \max_{m \in [Lmin,Lmax]}(E(m)). \qquad (6)$$

[0070]   In one embodiment, E(m) is the energy of a selected subband (e.g., LH or

HL or HH). In an alternative embodiment, E(m) is the energy of a sum of weighted

subbands [e.g. $\alpha E_{LH}(m) + \beta E_{HL}(m) + \gamma E_{HH}(m)$. In yet another embodiment, E(m) is the

maximum of energy from all subbands (e.g., $\max(E_{LH}(m), E_{HL}(m), E_{HH}(m))$.

[0071]   For energy, e.g., comparing formula (6) with (5), the importance measure

M in Eq. 1 measures the energy of wavelet coefficients at level m and is expressed as

$$Cost(M(W_m^{LL}(X)) = \alpha E(W_m^{HL}(X)) + \beta E(W_m^{LH}(X)) + \gamma E(W_m^{HH}(X)) \text{ or} \qquad (7)$$

$$Cost(M(W_m^{LL}(X)) = \max[E(W_m^{HL}(X)), E(W_m^{LH}(X)), E(W_m^{HH}(X))]. \qquad (8)$$

In case of entropy M measures entropy H of subbands,

$$Cost(M(W_m^{LL}(X)) = \alpha H(W_m^{HL}(X)) + \beta H(W_m^{LH}(X)) + \gamma H(W_m^{HH}(X)) \text{ or} \qquad (9)$$

$$Cost(M(W_m^{LL}(X)) = \max[H(W_m^{HL}(X)), H(W_m^{LH}(X)), H(W_m^{HH}(X))]. \qquad (10)$$

Otherwise computations are the same.

[0072]   Figure 9 is a flow diagram of one embodiment of a process for computing

an image representation that is visually recognizable when being displayed on a given

display device according to one embodiment. Referring to Figure 9, processing logic

receives an image (processing block 901). Processing logic then creates a smaller

representation of the image from a wavelet representation of the image, where the size of

the smaller representation of the image is selected based on the content of the image and

physical properties of a display device to display the smaller representation of the image

(processing block 902). As shown in Figure 9, as part of creating the smaller

20

representation, processing logic downsamples the image a number of times (processing block 903). The number of times is a number of times at which a trend changes from an importance measure increasing each time to the importance measure decreasing each time, such as shown, for example, in Figure 30.

[0073]   Figure 10 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 10, processing logic receives an image (processing block 1001). Processing logic then creates a smaller representation of the image from a wavelet representation of the image where, the size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 1002). As shown in Figure 10, as part of creating the smaller representation, processing logic downsamples the image a number of times (processing block 1003). The number of times is a number of times at which a trend changes from an importance measure increasing each time to the importance measure decreasing each time, such as shown, for example, in Figure 30. The importance measure is energy. Therefore, the number of times is a number of times at which a trend changes from an energy increasing each time to the energy decreasing each time.

[0074]   Figure 11 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 11, processing logic receives an image (processing block 1101). Processing logic then creates a smaller representation of the image from a wavelet representation of the image, where the size of the smaller

representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 1102). As shown in Figure 11, as part of creating the smaller representation, processing logic downsamples the image a number of times (processing block 1103). The number of times is a number of times at which a trend changes from an importance measure increasing each time to the importance measure decreasing each time, such as shown, for example, in Figure 30. In this case, the importance measure is an importance measure of wavelet coefficients in a selected subband. Therefore, the number of times is a number of times at which a trend changes from an importance measure of wavelet coefficients in a selected subband increasing each time to the importance measure of wavelet coefficients in a selected subband decreasing each time.

[0075]   Figure 12 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 12, processing logic receives an image (processing block 1201). Processing logic then creates a smaller representation of the image from a wavelet representation of the image, where the size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 1202). As shown in Figure 4, as part of creating the smaller representation, processing logic downsamples the image a number of times (processing block 1203). The number of times is a number of times at which a trend changes from an importance measure increasing each time to the importance measure decreasing each time, such as shown, for example, in Figure 30. In this case, the importance measure is

an importance measure of a sum of wavelet coefficients in weighted subbands.
Therefore, the number of times is a number of times at which a trend changes from an importance measure of a sum of wavelet coefficients in weighted subbands increasing each time to the importance measure of a sum of wavelet coefficients in weighted subbands decreasing each time.

[0076]   Figure 13 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 13, processing logic receives an image (processing block 1301). Processing logic then creates a smaller representation of the image from a wavelet representation of the image where the size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 1302). As shown in Figure 4, as part of creating the smaller representation, processing logic downsamples the image a number of times (processing block 1303). The number of times is the times at which a trend changes from an importance measure increasing each time to the importance measure decreasing each time, such as shown, for example, in Figure 30. In this case, the importance measure is a maximum importance measure of wavelet coefficients from all subbands. Therefore, the number of times is the times at which a trend changes from a maximum importance measure of wavelet coefficients from all subbands increasing each time to the maximum importance measure of wavelet coefficients from all subbands decreasing each time.

[0077]   In images that contain a lot of noise, the determination of $L_{display}$ could result in a scale containing mostly noise. Therefore, an additional criterion for

distinguishing between noise and the image is useful to eliminate this case. In one

embodiment, noise is removed from the wavelet coefficients using wavelet denoising.

One denoising technique that may be used for this is described in Donoho, D.L.,

"Denoising by soft-thresholding," IEEE Transactions on Information Theory, vol. 41, no.

3, pp. 613-627, 1995). Alternatively, another technique that may be used is described in

U.S. Patent Application Serial Number _____, entitled "Header-Based

Processing Of Images Compressed Using Multi-Scale Transforms", concurrently filed

January 10, 2002, and assigned to the corporate assignee of the present invention. In one

embodiment, after denoising, the determination of the display scale is performed as set

forth in Eq. (6) above.

[0078]   Figure 14 is a flow diagram of one embodiment of a process for

computing an image representation that is visually recognizable when being displayed on

a given display device. Referring to Figure 14, processing logic receives an image

(processing block 1401). Processing logic then creates a smaller representation of the

image from a wavelet representation of the image, where the size of the smaller

representation of the image is selected based on the content of the image and physical

properties of a display device to display the smaller representation of the image

(processing block 1402). As shown in Figure 14, as part of creating the smaller

representation, processing logic denoises wavelet coefficients (processing block 1403).

[0079]   In an alternative embodiment, in order to lower computational complexity,

wavelet coefficients are not considered at all possible scales (levels) simultaneously, but

are considered starting with the coarsest scale first, incorporating the next finer scales one

by one and checking after incorporating each level whether a display scale $L_{display}$ can be determined. One embodiment of such an approach is described below.

[0080]   In one approach, processing logic determines the energy as a function of level E(m). For increasing m, as long as $E(m-1) \leq E(m)$, processing logic computes E(m+1). If $m = L_{min}$, then processing logic sets $L_{display} = L_{min}$.

[0081]   If $E(m-1) > E(m)$, then processing level neglects level m as the possible display scale $L_{display}$ if the ratio of error between denoised and original coefficients in one or more subbands at level m-1 and the original coefficients is larger than a predetermined threshold. For example, in one embodiment, the predetermined threshold is 0.7. Otherwise, processing logic chooses $L_{display} = m-1$.

[0082]   Figure 15 illustrates results of an approach for computing an image representation that is visually recognizable when being displayed on a given display device for two example images.

[0083]   In an alternative embodiment, instead of determining the trend based on the importance measure energy, the importance measure entropy could be used. The entropy determines how many bits are spent to encode coefficients. The display scale is determined in a similar fashion as described in Eq. (8) by substituting entropy for energy of coefficients at level m.

[0084]   Figure 16 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 16, processing logic receives an image (processing block 1601). Processing logic then creates a smaller representation of the image from a wavelet representation of the image where the size of the smaller

representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 1602). As shown in Figure 16, as part of creating the smaller representation, processing logic downsamples the image a number of times (processing block 1603). The number of times is a number of times at which a trend changes from an importance measure increasing each time to the importance measure decreasing each time where the importance measure is entopy. Therefore, the number of times is a number of times at which a trend changes from entropy increasing each time to entropy decreasing each time.

[0085]   In one embodiment, some alternatives for energy include sum of entropy in all subbands, using the maximum energy, and using linear combination, such as described above.

SCALE SELECTION BY MAXIMAL SIGNIFICANCE RATIO

[0086]   In an alternative embodiment, a different criterion than that of Eq. (6) is used to determine a display scale. This criteria is based on a two-class labeling in which all wavelet coefficients are segregated into classes of significant and insignificant coefficients and then the scale that has the highest ratio of significant to insignificant coefficients is determined such that

$$L_{display} = \arg \max_{m=Lmin,...,Lmax}(\mu(\text{significant coefficients})/\mu(\text{insignificant coefficients})) \quad (12)$$

where μ measures a class of data. In one embodiment, this measure is a count of data points in the class. In an alternative embodiment, this measure is a weighted counting (e.g., weighting the center of the image more).

26

[0087]   In two-class labeling, M measures percentage of significant wavelet

coefficients at level m. The significance can be determined by checking whether a

coefficient has magnitude larger than a predefined threshold or not. Another method

would be to incorporate information from scales larger than m in the decision. One

example is to check whether a coefficient is larger than a given percentage of the energy

of coefficients at the next larger scale. The percentage threshold could depend on the

display device or the application. Another example is to estimate a local entropy for each

coefficient and check whether that value is larger than a given percentage of the entropy

of coefficients at the next larger scale.

[0088]   In one embodiment, the cost function in Eq. 1 is as follows:

$Cost$(M(W_m^{LL}(X)) = 0 if percentage of significant coefficients at level m is

smaller than a given threshold T, or given that percentage of significant coefficients at

level n ≥ m is smaller than T.

$Cost$(M(W_m^{LL}(X)) = percentage of significant coefficients at level m if that is

larger or equal than a given threshold T, given that percentage of significant coefficients

at level n ≥ m is larger than T.

[0089]   .Figure 17 is a flow diagram of one embodiment of a process for

computing an image representation that is visually recognizable when being displayed on

a given display device. Referring to Figure 17, processing logic receives an image

(processing block 1701). Processing logic then creates a smaller representation of the

image from a wavelet representation of the image, where the size of the smaller

representation of the image is selected based on the content of the image and physical

properties of a display device to display the smaller representation of the image

(processing block 1702). As shown in Figure 17, as part of creating the smaller representation, processing logic downsamples the image a number of times (processing logic 1703). The number of times is a number of times at which a ratio of wavelet coefficients in a class of significant coefficients to coefficients in a class of insignificant coefficients is highest.

[0090] Approaches to solving the two-class labeling problem are well known in the field of pattern recognition. For example, see Duda, R.O., Hart, P.E., and Stork, D.G., *Pattern Classification* (2nd ed.), Wiley, New York, 2000. The two-class labeling problem for each scale solved may be solved using Bayesian decision theory. Image data that can enter into the probability models are magnitude of coefficients or local entropies.

LOCAL SCALE SELECTION

[0091] Local scale selection needs a partition in the image or wavelet domain and needs to consider groupings of coefficients or pixels. The grouping can be either given by the application or the user or is determined from the image data.

PARTITION OF IMAGE DOMAIN INTO SEGMENTS

[0092] In one embodiment, the image is divided into two dimensional segments (e.g., tiles in J2K) and then global scale selection is performed as described below.

[0093] In one embodiment, instead of selecting a display scale and then choosing the LL component at that scale as the image representation, a part of the image at a specific scale (e.g., text at fine scale, background at coarse scale) is selected. In order to perform such a local scale selection, the image is partitioned into segments. In one embodiment, the segments are individual coefficients. In an alternative embodiment, the

segments are cover groups of coefficients, shaped like, for example, squares, rectangles, etc.

[0094] After a partition (given, e.g., by the segment size and shape size) is chosen, the same approach on global scale selection as described above is applied to each of the segments S of the partition. As a result, each S(i) has an assigned display scale $L_{display}(i)$. In terms of pattern recognition, the result is an $(L_{max}-L_{min})$-class labeling of the segment.

[0095] Figure 18 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 18, processing logic receives an image (processing block 1801). Processing logic then creates a smaller representation of the image from a wavelet representation of the image, where the size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 1802). As part of the creation process, processing logic partitions the image into segments (processing block 803). Processing logic then downsamples each segment a number of times (processing block 1804). Downsampling a cell is performed in the same manner as is downsampling an image.

[0096] In one embodiment, only parts (e.g., segments) of the image that are associated with a particular display scale (e.g., $L_{display} = 2$) are displayed. The selected segments may be grouped or embedded into connected components. For example, for most image displays a rectangular display of image information is preferred because most image displays are rectangular. There are many approaches to selecting rectangles

depending on an m-class labeling result. In one embodiment, the largest bounding box that contains all segments with $L_{display} = 2$ is chosen. The disadvantage of this embodiment is that isolated outliers for the box might be very large. According to an alternative embodiment, an approach that avoids this effect is used by choosing the rectangle such that the ratio

$$R(2) = \mu(\text{segments in rectangle with } L_{display}=2)/\mu(\text{segments in rectangle with } L_{display} \neq 2) \quad (13)$$

is maximized, where μ denotes a measure as in Eq. (12). An example of one such rectangle is given in Figure 10B. In an alternative embodiment, an approach is used that selects the rectangle that maximizes

$$R(m) = \mu(\text{segments in rectangle with } L_{display}= m)/m(\text{segments in rectangle with } L_{display} \neq m) \quad (14)$$

over all $m \in \lambda$.

[0097]   Differently stated, $(L_{max}-L_{min})$-class labeling problems combining global and local scale selections lead to design choices for reduced images in specific applications that are discussed in further detail below.

PARTITION OF WAVELET DOMAIN INTO CELLS

[0098]   When partitioning the wavelet domain into cells, the wavelet domain is divided into wavelet cells. The metric M(x) measures information content for a group of data points a low-resolution X_m.

[0099]   Let $x \in X_m$. The information content of x is computed by M(x) from measuring the contribution of energy or entropy of all the wavelet cells level larger or equal M that include x in their support at resolution level m. The energy or entropy may be scaled depending on the resolution of the wavelet cell and its size. The cost function sums over the measured information content of all data points that fit into a predescribed

30

shape (e.g., a rectangle) or a given size. By choosing those points that produce the largest output of the cost function a part of the original image at a specific resolution is chosen for a reduced size image with recognizable information.

[0100]   Figure 19 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 19, processing logic receives an image (processing block 1901). Processing logic then creates a smaller representation of the image from a wavelet representation of the image, where the size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 1902). As part of the creation process, processing logic partitions the image into cells using, for example, JPEG2000 code units (processing block 1903). Thereafter, processing logic displays image content described by coefficients in selected cells, where cells are selected through a segmentation algorithm. For more information see co-pending U.S. Patent Application Serial No. _____, entitled *Header-Based Processing Of Images Compressed Using Multi-Scale Transforms*, concurrently filed on January 10, 2002, and assigned to the corporate assignee of the present invention.

[0101]   Figure 20 is a flow diagram of one embodiment of a process for creating a reduced size image based on global scale selection. In process block 2001, image data is received. In process block 2002, data describing the importance of wavelet coefficients in an L-level wavelet decomposition is computed/extracted. The data should be level dependent and could be the energy or entropy of wavelet coefficients at given levels of resolution. In process block 2003, the display scale $L_{display}$ is determined. If $L_{display} <$

31

L_{min} then $L_{display}$ is set to equal L_{min}. If $L_{display} > L_{max}$ then $L_{display}$ is set to equal L_{max}. Finally, in process block 2004, the LL component at level $L_{display}$ of wavelet decomposition is displayed.

[0102]   Figure 21 is a flow diagram of one embodiment of a process for creating a reduced size image based on local scale selection. In process block 2101, image data, application specific design choices, and user preferences for output display are received. In process block 2102, a partition of an L-level wavelet decomposition into cells C is chosen. In process block 2103, data describing the importance of wavelet coefficients in an L-level wavelet decomposition is computed/extracted. The data should be level dependent and could be the energy or entropy of wavelet coefficients at given levels of resolution. In process block 2104, the display scale $L_{display}(i)$ is determined for each cell C(i). If $L_{display}(i) < L_{min}$ then $L_{display}(i)$ is set to equal to L_{min}. If $L_{display}(i) > L_{max}$ then $L_{display}(i)$ is set to equal L_{max}. In process block 2105, classes of cells with the same display scales (L_{min}-L_{max} classes) are formed. In process block 2106, application specific design choices such as shape and size are considered and the best fit of groups of cells belonging to the same display scale class in a given design in determined. Finally, in process block 2107, pure or modified LL components of selected grouping at their class' display level are displayed.

[0103]   Figure 22 is a flow diagram of one embodiment of a process for creating a reduced size image based on a combination of global and local scale selection. In process block 2201, image data, application specific design choices, and user preferences for output display are received. In process block 2102, data describing the importance of wavelet coefficients in an L-LEVEL wavelet decomposition is computed/extracted. The

32

data should be level and cell dependent and could be the energy or entropy of wavelet coefficients at given levels of resolution. In process block 2203, display scale $L_{DISPLAY}$ is determined. If $L_{DISPLAY} \notin [L_{MIN}, L_{MAX}]$ in decision diamond 2204, then process block 2205 performs local scale selection on the image such that selected LL components have display level $L \notin [0, L(MAX)]$ and fit into display pixel size or fixed size and shape. Otherwise, in process block 2207, the LL component at level $L_{DISPLAY}$ of wavelet decomposition is displayed (i.e., global scale decomposition).

[0104]   As an extension the maximization could be computed over a pre-described shape, but variable size or flexible shape of fixed size or flexible shapes of variable size.

[0105]   Below is a mathematical formulation of the above.

A partition P of a set X is a set of subsets of X such that $P = \bigcup\limits_{p \in P} p$ and

$p_1 \cap p_2 = 0$ for $p_1 \neq p_2$ and every pair of sets $p_1, p_2 \in P$.

$$I_m^{HL} = \left\{ (i,j) \middle| i = 0,...M2^{-m} - 1, j = N2^{-m},...N2^{-2m} - 1 \right\} \qquad (15)$$

$$I_m^{LH} = \left\{ (i,j) \middle| i = M2^{-m},...M2^{-2m} - 1, j = 0,...N2^{-2m} - 1 \right\} \qquad (16)$$

$$I_m^{HH} = \left\{ (i,j) \middle| i = M2^{-m},...M2^{-2m} - 1, j = N2^{-m},...N2^{-2m} - 1 \right\} \qquad (17)$$

Let P_m^{HL}, P_m^{LH} and P_m^{HH} be partitions of I_m^{HL}, I_m^{LH} and I_m^{HH}, respectively.

[0106]   A reduced size image with recognizable content is given by an output border and output image content. The border is given by a shape (e.g., rectangle) and a size (e.g., 68x80 pixels). In order to fill the border with image content, a resolution of the image and an anchor point position to locate the border is determined (e.g., coordinate (10,20) in resolution image X_m). An anchor point for a rectangle or square could be the upper left corner, while the anchor point for a circle may be the center.

[0107] In a first case, given a predetermined shape and size of the output border, the maximization problem is stated as

$$(L_{display}, A_{display}) = \arg\max_{m \in \lambda, a \in A} \{Cost(M(X_m|_{output\_border})), shape(output\_border)$$
$$= fixed, size(output\_border) = fixed, anchor(output\_border) = variable\}. \quad (18)$$

The reduced size image with recognizable content is then defined as

$$\{x \in X_{Ldisplay}|_{output\_border}, shape(output\_border) = fixed, size(output\_border) = fixed, anchor$$
$$(output\_border) = A_{display}\}.$$

[0108] In a second case, in addition, if the size of the template of the reduced size image with recognizable content is variable the maximization problem is

$$(L_{display}, A_{display}, N_{display}) = \arg\max_{m \in \lambda, a \in A, n \in N} \{Cost(M(X_m|_{output\_border})), shape(output\_border)$$
$$= fixed, n = size(output\_border) = variable, a = anchor(output\_border) = variable\}. \quad (19)$$

The reduced size image with recognizable content is then defined as

$$\{x \in X_{Ldisplay}|_{output\_border}, shape(output\_border) = fixed, size(output\_border) = N_{display}, anchor(output\_border)$$
$$= A_{display}\}.$$

[0109] In a third case, if anchor point of the output border, the size and shape of the template of the reduced size image with recognizable content are variable the maximization problem is stated above.

$$(L_{display}, A_{display}, N_{display}, S_{display}) = \arg\max_{m \in \lambda, a \in A, n \in N, s \in S} \{Cost(M(X_m|_{output\_border})), s = shape(output\_border)$$
$$= variable, n = size(output\_border) = variable, a = anchor(output\_border) = variable\}. \quad (20)$$

The reduced size image with recognizable content is then defined as

$$\{x \in X_{Ldisplay}|_{output\_border}, shape(output\_border) = S_{display}, size(output\_border) =$$
$$N_{display}, anchor(output\_border) = A_{display}\}.$$

[0110]   In this case, the vector ($L_{display}$, $A_{display}$, $N_{display}$, $S_{display}$,) determines the reduced size image with recognizable content.

[0111]   In the first case, the location and resolution selection for the fixed reduced size image with recognizable content are image content and display device dependent. The size and shape of the reduced size image with recognizable content border are application or user dependent. In the second case, location, resolution and reduced size image with recognizable content size, and in the third case location, resolution size and shape of the reduced size image with recognizable content are image content and display device dependent.

[0112]   The four-dimensional vector ($L_{display}$, $A_{display}$, $N_{display}$, $S_{display}$) determines the reduced size image with recognizable content. An extension is possible by allowing the entries in the vector to be vectors. In this case a reduced size image with recognizable content would be a collection of selected resolution segments with specific output borders. The segments may be chosen such that their full-resolution versions cover the entire image, either overlapping or non-overlapping. This is referred to herein as a Multiscale Collage.

[0113]   In general, given an image the method creates an optimal thumbnails for a given set of constrains on the vector ($L_{display}$, $A_{display}$, $N_{display}$, $S_{display}$). In this case, a reduced size image with recognizable content is a combination of specifically selected resolution segments. In the scalar and vector case, postprocessing can be applied to the resolution segments, such as interpolation, extrapolation, warping, blurring, etc.

SCALE SELECTION USING META DATA CREATED BY A JPEG2000 IMAGE CODER

[0114]   The codestream of an image created by a JPEG2000 ("J2K") image coder orders the wavelet domain image data into organizational units, such as packets, codeblocks, or tiles. Those units contains header information including features such as a number of zero bitplanes or unit code length. In one embodiment, those features are used to determine the importance measure M (e.g., entropy) and the cost function Cost. The partition of wavelet domain image data is then given by the choice of the code units (e.g., code blocks). Those data are easily accessible without decoding the actual image data and lead to very fast implementations of reduced size image creations.

[0115]   In order to perform scale selection using J2K meta data, the partitioning of the wavelet domain into cells is given by codeblocks. The image domain may be given by tiles.

[0116]   For global scale section, the entropy of a subband is given as the sum of length of codeblocks contained in the subband. The metrics described above in conjunction with global scale selection may be applied and the reduced size image with recognizable content computed.

[0117]   For local scale selection, if the image is given in tiles, then the local scale selection approach described above with respect to partitioning the image domain into segments may be used. Alternatively, for local scale selection using J2K meta data, the local scale selection approach described above with respect to partitioning the wavelet domain into cells may be used where the wavelet cells are the code blocks and M measures the entropy of a codeblock.

[0118]   A combination of global and local scale section may be used. For example, if the output border size and shape are not constrained, global selection may be

36

attempted initially. After doing so, if the calculated resolution is out of the range x, then local scale selection may be started. In one embodiment, local scale section is started under a constraint (e.g., a rectangular shape). For example, if image is square, then global scale selection produces a square. If global scale selection does not work, then local scale selection should be applied given the constraint that the shape should be a square not bigger than a specific size.

[0119]   The following is an example for combined global and local scale selection for two different displays. Two different display devices are considered, one is a CRT desktop monitor and one an LCD display of a PDA device. The physical properties of those displays are listed in Table 1.

Table 1 - Physical characteristics of two different display devices

	absolute pixel resolution	dpi resolution	contrast ratio
CRT desktop	1280 x 1024	95	260 :1
LCD PDA	320 x 320	140	190:1

[0120]   Table 2 lists the calculated parameters L_{min}, D_{min} and L_{max}. The parameter C_{device} is set, e.g., by tuning or by using the ratio of the display contrast ratio to the contrast ratio of a reference display.

Table 2 - Parameters derived from display properties

	CRT desktop	LCD PDA
L_{min}	0	2
D_{min}	7	11
C_{device}	1	260/190 = 1.37
L_{max}	7	4

37

[0121]   Global scale selection determines $L_{display} = 5$ for the CRT device. For the

LCD device global scale selection is $L_{display} = 5$. But that is larger than L_{max}. Therefore local

scale selection starts, a square shape of 300 x 300 pixels is fixed and it is searches for the

location and resolution L in the interval $[0, L_{max}]$ that results in the best fit, i.e. maximizes

the cost function.

APPLICATIONS

[0122]   There are several ways to incorporate the concept of global and local scale

selection into possible design choices for creating reduced size image. In general, they

can be separated into five groups.

[0123]   In one embodiment, an image is created that contains information of only

one image segment such that the LL component of this segment at its calculated display

scale fits into a given shape of fixed size (e.g., a 68x88 pixel rectangle). In this manner,

the image then contains the LL component of the segment at a calculated display scale.

[0124]   Figure 23 is a flow diagram of one embodiment of a process for

computing an image representation that is visually recognizable when being displayed on

a given display device. Referring to Figure 23, processing logic receives an image

(processing block 2301). Processing logic then creates a smaller representation of the

image from a wavelet representation of the image, where the size of the smaller

representation of the image is selected based on the content of the image and physical

properties of a display device to display the smaller representation of the image

(processing block 2302). As shown in Figure 23, as part of creating the smaller

representation, processing logic downsamples the image a number of times (processing

38

block 2303). The number of times is a number of times at which a size of a selected segment best approximates a given fixed size.

[0125]   In one embodiment, a reduced size image is created that contains information of only one segment of the image at a particular LL resolution level $L_{display}$. The shape of the reduced size image is fixed, but its size is flexible. For example, a reduced size image shape could be a rectangle of a certain width/height ratio R. The reduced size image then contains the image segment at its associated resolution, such that it satisfies the "best-fit" criterion over all considered rectangles of ratio R. .

[0126]   Figure 24 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on a given display device. Referring to Figure 24, processing logic receives an image (processing block 2401). Processing logic then creates a smaller representation of the image from a wavelet representation of the image, where the size of the smaller representation of the image is selected based on the content of the image and physical properties of a display device to display the smaller representation of the image (processing block 2402). As shown in Figure 24, as part of creating the smaller representation, processing logic downsamples the image a number of times (processing block 2402). The number of times is a number of times at which a shape of the segment most approximates a given fixed shape.

[0127]   In one embodiment, a reduced size image is created that contains information of only one segment of the image at a particular resolution. The reduced size image shape and size are then flexible.

[0128]   Figure 25 is a flow diagram of one embodiment of a process for computing an image representation that is visually recognizable when being displayed on

39

a given display device. Referring to Figure 25, processing logic receives an image

(processing block 2501). Processing logic then creates a smaller representation of the

image from a wavelet representation of the image, where the size of the smaller

representation of the image is selected based on the content of the image and physical

properties of a display device to display the smaller representation of the image

(processing block 2402). As shown in Figure 25, as part of creating the smaller

representation processing logic downsamples the image a number of times (processing

block 2503). The number of times is a number of times at which a resolution of the

selected segment best approximates a given fixed resolution.

[0129]   In one embodiment, a reduced size image is created that contains several

segments of the image, each displayed at the LL component at its display level (i.e., a

multiscale collage).

[0130]   In one embodiment, a reduced size image is created that contains several

segments of the image. Each segment is displayed at the LL component at its display

level, with post-processing further performed on the LL component that affects the values

of LL coefficients as well as the shapes of the displayed segments. Examples of such

post-processing operations include extrapolation, blurring, stylizing, enhancing, fusing,

warping, morphing, etc. The design space for reduced size images is symbolized in

Table 3.

Table 3 - Matrix Of Reduced Size Image Design Features

	Fixed(f)	selection from list(s)	unconstrained(u)
Shape	Prior art	New invention	New invention
Size	Prior art	New invention	New invention
Location	Prior art	New invention	New invention
Resolution	Prior art	New invention	New invention
post-processing steps, e.g., blurring, warping.	New invention	New invention	New invention

[0131]   Figure 26 illustrates reduced size image design choices according to one

embodiment. In one embodiment, from entire image 2601, an image segment 2602 is

selected. A smaller representation 2603 of the segment 2602 is created by downsampling

the segment 2602 a number of times. The number of times is a number of times at which

a size of the segment most approximates a given fixed size (e.g., 68 x 88 pixels). In an

alternative embodiment, from entire image 2604, an image segment 2605 is selected. A

smaller representation 2606 of the segment 2605 is created by downsampling the segment

2605 a number of times. The number of times is a number of times at which a resolution

of the segment best approximates a given fixed resolution (e.g., a rectangle of a fixed

width/height ratio). In an alternative embodiment, from entire image 2607, an image

segment 2608 is selected. A smaller representation 2609 of the segment 2608 is created

by downsampling the segment 2608 a number of times. The number of times is a number

of times at which a shape of the segment best approximates a given fixed shape (e.g., a

shape selected from a list of shapes and flexible sizes).

[0132]   Figure 27 illustrates exemplary reduced size image design choices. In one

embodiment, selected image segments 2701-2707 are each downsampled to varying

extents. The resulting multiscale collage contains, for example, segments 2708, 2712,

and 2713 at 25% of the sizes of their original counterparts. Segments 2702 and 2703 are,

41

for example, 50% of the sizes of their original counterparts. Segment 2714 is at just 12.5% of its original counterpart. Finally, segment 2711 remains at 100% of the size of its original counterpart, segment 2704.

[0133]   In one embodiment, selected image segments 2715-2721 are upsampled from an LL component by a factor. In the resulting multiscale collage, for example, segments 2722, 2726, and 2727 are upsampled by a factor of 2x from their scale selected LL components. Segments 2723 and 2724 are upsampled by a factor of 1x from their scale selected LL components. Segment 2728 is upsampled by a factor of 3x from its scale selected LL components. Finally, segment 2725 remains at the original size from its original counterpart, segment 2718.

[0134]   In one embodiment, selected image segments 2729-2735 are upsampled in the same manner as segments 2715-2721. Furthermore, each corresponding segment 2736-2742 in a resulting multiscale collage is post-processing enhanced, warped, blurred, etc.

[0135]   The embodiments above may be applied to digital camera displays, digital photo album software, image browsing, video summaries, page layout, and web page design. In applications such as these, the display range parameters L_{min} and L_{max} are in one embodiment set by the user and are potentially adjusted during the use of the reduced size image creation system. In one embodiment operating system, parameters such as absolute pixel resolution are extracted from the display settings without requiring any user effort. In the above description, only downsampling by a factor of two was considered due to its natural appearance in the wavelet decomposition. It is clear to one skilled in the art, however, that the approaches to reduced size image creation can also be

extended to sampling ratios other than powers of two by interpolation or extrapolation of selected segments at specific resolutions.

[0136] Figure 28 is a schematic diagram illustrating a device for computing an image representation that is visually recognizable when being displayed on a given display device according to one embodiment. The device 2801 comprises a receiving unit 2802 to receive an image, and a processing unit 2803 coupled with the receiving unit 2802. The processing unit 2803 creates a smaller representation of the image from a wavelet representation of the image, wherein a size of the smaller representation of the image is selected to compensate for a content of the image and physical properties of a display device (not shown) to display the smaller representation of the image. In one embodiment, to create a smaller representation of the image from a wavelet representation of the image, processing unit 2803 downsamples the image a number of times, wherein the number of times is large enough to cause the smaller representation of the image to be completely visible on the display device. In one embodiment, to create a smaller representation of the image from a wavelet representation of the image, processing unit 2803 downsamples the image a number of times, wherein the number of times depends proportionately on a ratio of coarse structures in the image to fine structures in the image. In one embodiment, to create a smaller representation of the image from a wavelet representation of the image, processing unit 2803 denoises coefficients. In one embodiment, to create a smaller representation of the image from a wavelet representation of the image, processing unit 2803 downsamples the image a number of times, wherein the number of times is a number of times at which a ratio of wavelet coefficients in a class of significant features to wavelet coefficients in a class of insignificant features is highest. In one embodiment, to create a smaller representation of

the image from a wavelet representation of the image, processing unit 2803 downsamples a segment of the image a number of times, wherein the number of times is a number of times at which a size of the segment most approximates a given fixed size. In one embodiment, to create a smaller representation of the image from a wavelet representation of the image, processing unit 2803 downsamples a segment of the image a number of times, wherein the number of times is a number of times at which a shape of the segment most approximates a given fixed shape. In one embodiment, to create a smaller representation of the image from a wavelet representation of the image, processing unit 2803 downsamples a segment of the image a number of times, wherein the number of times is a number of times at which a resolution of the segment most approximates a given fixed resolution. In one embodiment, to create a smaller representation of the image from a wavelet representation of the image, processing unit 2803 downsamples the image a number of times, wherein the number of times is small enough to cause a number of dots in a diameter of the object to be at least as large as a number of dots in the minimal visible object diameter. In one embodiment, to create a smaller representation of the image from a wavelet representation of the image, processing unit 2803 downsamples the image a number of times, wherein the number of times is a number of times at which a trend changes from an importance measure increasing each time to the importance measure decreasing each time. In one embodiment, to create a smaller representation of the image from a wavelet representation of the image, processing unit 2803 partitions the image into cells segments and downsample segments. In one embodiment, to partition the image into cells, processing unit 2803 partitions the image by JPEG2000 code units.

44

[0137]   The techniques described herein may be extended to color. For color images, all these techniques can be used on one chromatic band (e.g., L* inCIEL*uv, Yin Yuv, or G in sRGB) or on some or all. If more than one chromatic band is used, a technique can be used to combine results. Also, the techniques described herein may be extended to time (e.g., motion, video). In such a case, objects may be identified to track in thumbnail and/or key frame may be located, and then subsequently processed.

An Exemplary Computer System

[0138]   Figure 29 is a block diagram of an exemplary computer system that may perform one or more of the operations described herein. Referring to Figure 29, computer system 2900 may comprise an exemplary client 2950 or server 2900 computer system. Computer system 2900 comprises a communication mechanism or bus 2911 for communicating information, and a processor 2912 coupled with bus 2911 for processing information. Processor 2912 includes a microprocessor, but is not limited to a microprocessor, such as, for example, Pentium™, PowerPC™, etc.

[0139]   System 2900 further comprises a random access memory (RAM), or other dynamic storage device 2904 (referred to as main memory) coupled to bus 2911 for storing information and instructions to be executed by processor 2912. Main memory 2904 also may be used for storing temporary variables or other intermediate information during execution of instructions by processor 2912.

[0140]   Computer system 2900 also comprises a read only memory (ROM) and/or other static storage device 2906 coupled to bus 2911 for storing static information and instructions for processor 2912, and a data storage device 2907, such as a magnetic disk

or optical disk and its corresponding disk drive. Data storage device 2907 is coupled to bus 2911 for storing information and instructions.

[0141]   Computer system 2900 may further be coupled to a display device 2921, such as a cathode ray tube (CRT) or liquid crystal display (LCD), coupled to bus 2911 for displaying information to a computer user. An alphanumeric input device 2922, including alphanumeric and other keys, may also be coupled to bus 2911 for communicating information and command selections to processor 2912. An additional user input device is cursor control 2923, such as a mouse, trackball, trackpad, stylus, or cursor direction keys, coupled to bus 2911 for communicating direction information and command selections to processor 2912, and for controlling cursor movement on display 2921.

[0142]   Another device that may be coupled to bus 2911 is hard copy device 2924, which may be used for printing instructions, data, or other information on a medium such as paper, film, or similar types of media. Furthermore, a sound recording and playback device, such as a speaker and/or microphone may optionally be coupled to bus 2911 for audio interfacing with computer system 2900. Another device that may be coupled to bus 2911 is a wired/wireless communication capability 2925 to communication to a phone or handheld palm device.

[0143]   Note that any or all of the components of system 2900 and associated hardware may be used in the present invention. However, it can be appreciated that other configurations of the computer system may include some or all of the devices.

[0144]   Whereas many alterations and modifications of the present invention will no doubt become apparent to a person of ordinary skill in the art after having read the foregoing description, it is to be understood that any particular embodiment shown and

described by way of illustration is in no way intended to be considered limiting.

Therefore, references to details of various embodiments are not intended to limit the

scope of the claims which in themselves recite only those features regarded as essential to

the invention.